(8)

NAME_____

PERIOD_____ DATE_____

THE VALUE OF r
Correlations

The points in a scatter plot can have a positive, negative or no correlation.

Positive correlation = the points can be approximated by a line that has a positive slope.

Negative correlation = the points can be approximated by a line that has a negative slope.

No correlation = there is no correlation which can be found between the x and y values.

In statistics, the value of r is called the correlation coefficient and is used as a measure of how well a collection of data can be modeled by a line. Correlation coefficients range between -1 and 1. The closer |r| is to 1, the better the line fits the points.

The formula for r is

$$r = \frac{\text{Sum of } (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\left[\text{Sum of } (x_i - \bar{x})^2\right]\left[\text{Sum of } (y_i - \bar{y})^2\right]}}$$

where $\bar{x}$ is the average x-value and $\bar{y}$ is the average y-value.

We will simplify this formula to

$$r = \frac{A}{\sqrt{B \cdot C}}$$

Correlation Coefficient - example

	x	$x_i - \bar{x}$	$(x_i - \bar{x})^2$	y	$y_i - \bar{y}$	$(y_i - \bar{y})^2$	$(x_i - \bar{x})(y_i - \bar{y})$
	0			12			
	10			22			
	20			30			
	30			38			
	40			54			
Sum		//////////////			//////////////		
Average			B			C	A

$$r = \frac{A}{\sqrt{B \cdot C}} =$$

Find the correlation coefficient for the following data.

	x	$x_i - \bar{x}$	$(x_i - \bar{x})^2$	y	$y_i - \bar{y}$	$(y_i - \bar{y})^2$	$(x_i - \bar{x})(y_i - \bar{y})$
	6			6			
	7			8			
	8			12			
	9			19			
	10			25			
Sum		////////			////////		
Average			B			C	A

$$r = \frac{A}{\sqrt{B \cdot C}} =$$